Convexification of the Range-Only Station Keeping Problem

Cao, Ming; Morse, A. Stephen

Published in:
Analysis and Design of Nonlinear Control Systems

Publication date:
2008

Convexification of the Range-Only Station Keeping Problem

Ming Cao and A. Stephen Morse

Department of Electrical Engineering, Yale University, New Haven, CT 06520, USA

This paper is dedicated to Alberto Isidori on the occasion of his 65th birthday.

Summary. Using concepts from switched adaptive control theory plus a special parameterization of the class of 2×2 nonsingular matrices, a tractable and provably correct solution is given to the three landmark station keeping problem in the plane in which range measurements are the only sensed signals upon which station keeping is to be based. The performance of the overall system degrades gracefully in the face of increasing measurement and miss-alignment errors, provided the measurement errors are not too large.

1 Introduction

"Station keeping" is a term from orbital mechanics which refers to the "practice of maintaining the orbital position of satellites in geostationary orbit" {Wikipedia}. In this paper as in [3], we take station keeping to mean the practice of keeping a mobile autonomous agent in a position in the plane which is determined by prescribed distances from two or more landmarks. We refer to these landmarks as neighboring agents because we envision solutions to the station keeping problem as potential solutions to multi-agent formation maintenance problems. We are particularly interested in solutions to the station keeping problem in which the only signals available to the agent whose position is to be maintained, are noisy range measurements from its neighbors[1]. Our approach to station keeping builds on the work initiated in [3] where we treated station keeping as a problem in switched adaptive control. We continue with the same approach in this paper but now deal directly with an important computational issue which was not addressed in [3]. In particular, the control system considered in [3] requires an algorithm capable

[1] We are indebted to B. D. O. Anderson for making us aware of this problem.

of minimizing with respect to the four entries in a 2×2 nonsingular matrix P, a cost function of the form $M(X, P) = \text{trace}\{\begin{bmatrix} I & P \end{bmatrix} X \begin{bmatrix} I & P \end{bmatrix}'\}$ where X is a 4×4 positive semi-definite matrix. What makes the problem difficult is the constraint that P must be non-singular, since this leads to an non-convex optimization problem. The main contribution of this paper is to explain how to avoid this difficulty by utilizing the fact that any 2×2 non-singular matrix B can be written as $B = U(I + L)S$ where U is a specially structured matrix from a finite set, L is strictly lower triangular and S is symmetric and positive definite [12]. This fact enables us to modify the optimization problem just described, so that instead of having a non-convex problem to solve, one has a finite set of convex problems instead. Not only does the modification lead to convex programming problems, but also programming problems which can each be solved efficiently using semi-definite programming methods [22].

Work on the range-only station keeping problem already exists [9, 20, 2] and related work on range-only source localization can be found in [5, 4]. The station keeping problem is closely related to the Simultaneous Localization and Mapping {SLAM} problem [11, 16, 6, 23], which is also called the Concurrent Mapping and Localization problem [7, 21]. SLAM is the process of building a map of an unknown environment by using mobile robots' sensed information and simultaneously estimating those robots' locations by using this map. The station keeping problem with one autonomous agent and multiple landmarks can be cast as a SLAM problem in which the map describes the positions of the landmarks and the autonomous agent is the robot to be localized. There are several approaches to the SLAM problem, such as those based on Kalman filters [19, 1] and those using sequential Monte Carlo techniques [8, 24]. Kalman filtering based methods apply to linearized observation models and assume that the measurement errors are Gaussian. Since most of the sensory data from the range-only measurements are nonlinear and with non-Gaussian errors, the limitation of the Kalman filter method in this context is obvious. Sequential Monte Carlo based methods use nonlinear observation models and do not require suitable probabilistic models for measurement noises, but do require large numbers of samples; typically such methods are computationally difficult to implement. There are also several interesting and new set-based techniques addressed to the range-only SLAM problem [9, 20], but these have not been validated mathematically.

Several features of the station keeping method proposed here distinguish it from SLAM-based methods. First, SLAM algorithms seek to localize and map whereas the approach here focuses sharply and exclusively on the ultimate goal of moving an agent to its assigned position; no attempt is made to localize the assigned position and because of this, the approach taken here is fundamentally different than the more indirect SLAM approach. Second, the present method uses a provably correct switched adaptive control algorithm, whereas the SLAM-based methods do not.

In Section 2 we formulate the station keeping problem of interest. Error models appropriate to the solution to the problem are developed in Section 3. Some of the error equations developed have appeared previously in [9, 20, 18] and elsewhere. In Section 4 we present a switched adaptive control system which solves the three neighbor station keeping problem for a point modelled agent. The control system consists of a "multi-estimator" $\mathbb{E}$, a "multi-controller" $\mathbb{C}$, a "monitor" $\mathbb{M}$ and a "dwell-time switching logic" $\mathbb{S}$. These terms and definitions have been discussed before in [14, 15] and elsewhere. In Section 4.3, the output of the monitor is defined to be a parameter-dependent, scalar-valued signal of the form $M(W,P) = \text{trace}\{\begin{bmatrix} I & P\end{bmatrix} W \begin{bmatrix} I & P\end{bmatrix}'\}$ where W is a 4×4 positive semi-definite signal generated by the monitor and P is a 2×2 non-singular matrix of parameters taking values in a compact but non-convex parameter space $\mathcal{P}$. Although this particular definition is intuitive and natural for the adaptive solution to the station keeping problem, as we've already noted, the definition leads to non-convex optimization problem. To avoid this, use is made of the previously mentioned fact that any 2×2 non-singular matrix B can be written as $B = U(I + L)S$ where U is a specially structured matrix from a finite set, L is strictly lower triangular and S is symmetric and positive definite [12]. In Section 4.3 $M(\,\cdot\,)$ is redefined as $M(W,U,L,S) = \text{trace}\{\begin{bmatrix}(I-L)U' & S\end{bmatrix} X \begin{bmatrix}(I-L)U' & S\end{bmatrix}'\}$ where L and S take values in compact convex sets $\mathcal{L}$ and $\mathcal{S}$ respectively. More detailed descriptions of these sets are derived in Section 6. In Section 7 it is then explained how to re-formulate the resulting problem of minimizing $M(W,U,L,S)$ over $\mathcal{L} \times \mathcal{S}$ for fixed W and U, as a semi-definite programming problem.

Because of the re-parameterization just outlined, the resulting switched adaptive control is completely tractable and easy to implement. In addition, it has especially desirable properties. For example, in the absence of errors the control causes agent positioning to occur exponentially fast; moreover it guarantees that performance will degrade gracefully in the face of increasing measurement and miss-alignment errors, provided the measurement errors are not too large. In Section 5 we sketch the ideas upon which these claims are based.

2 Formulation

Let $n > 1$ be an integer. The system of interest consists of $n+1$ points in the plane labelled $0, 1, 2, \ldots, n$ which will be referred to as agents. Let $x_0, x_1, \ldots, x_n$ denote the coordinate vector of current positions of agents $0, 1, 2, \ldots n$ respectively with respect to a common frame of reference. Assume that the formation is suppose to come to rest and moreover that agents $1, 2, \ldots, n$ are already at their proper positions in the formation and are at rest. Thus

$$\dot{x}_i = 0, \quad i \in \{1, 2, 3, \ldots, n\}. \tag{1}$$

We further assume that the nominal model for how agent 0 moves is a kinematic point model of the form

$$\dot{x}_0 = u \tag{2}$$

where u is an open loop control taking values in $\mathbb{R}^2$.

Suppose that agent 0 can sense its distances $y_1, y_2, \ldots, y_n$ from neighboring agents $1, 2, \ldots, n$ with uniformly bounded, additive errors $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ respectively. Thus

$$y_i = ||x_i - x_0|| + \epsilon_i, \quad i \in \{1, 2, \ldots, n\} \tag{3}$$

where $||\cdot||$ denotes the Euclidean 2-norm. Suppose in addition that agent 0 is given a set of non-negative numbers $d_1, d_2, \ldots, d_n$, where d_i represents a desired distance from agent 0 to agent i. The problem is to devise a control law depending on the d_i and the y_i which, were the ϵ_i all zero, would cause agent 0 to move to a position in the formation which, for $i \in \{1, 2, \ldots, n\}$, is d_i units from agent i. We call this the *n neighbor station keeping problem.* We shall also require the controllers we devise to guarantee that errors between the y_i and their desired values eventually become small if the measurement errors are all small.

Let x^* denote the target position to which agent 0 would have to move were the station keeping problem solvable. Then x^* would have to satisfy

$$d_i = ||x_i - x^*||, \quad i \in \{1, 2, \ldots, n\}. \tag{4}$$

There are two cases to consider.

1) If $n = 2$, there will be two solutions x^* to (4) if $|d_1 - d_2| < ||x_1 - x_2|| < d_1 + d_2$ and no solutions if either $|d_1 - d_2| > ||x_1 - x_2||$ or $||x_1 - x_2|| > d_1 + d_2$. We will assume that two solutions exist and that the target position is the one closest to the initial position of agent zero.
2) If $n \geq 3$, there will exist a solution x^* to (4) only if agents 1 through n are aligned in such a way so that the circles centered at the x_i of radii d_i all intersect at least one point. If the x_i are so aligned and at least three x_i are not co-linear, then x^* is even unique. Such alignments are of course exceptional. To account for the more realistic situation when points are out of alignment, we will assume instead of (4), that there is a value of x^* for which
$$d_i = ||x^* - x_i|| + \bar{\epsilon}_i, \quad i \in \{1, 2, \ldots, n\} \tag{5}$$
where each $\bar{\epsilon}_i$ is a small miss-alignment error.

Our specific control objective can now be stated. Devise a feedback control for agent 0, using the d_i and measurements y_i, which bounds the induced $\mathcal{L}^2$ gains from each ϵ_i and each $\bar{\epsilon}_i$ to each of the errors

$$e_i = y_i^2 - d_i^2, \quad i \in \{1, 2, 3, \ldots, n\}. \tag{6}$$

We will address this problem using well known concepts and constructions from adaptive control.

3 Error Models

The controllers which we propose to study will all be based on suitably defined error models. We now proceed to develop these models.

3.1 Error Equations

To begin, we want to derive a useful expression for each e_i. In view of (3)

$$y_i^2 = ||x_i - x_0||^2 + 2\epsilon_i||x_i - x_0|| + \epsilon_i^2.$$

But

$$||x_i - x_0||^2 = ||x_i - x^*||^2 + 2(x^* - x_i)'\bar{x}_0 + ||\bar{x}_0||^2$$

where

$$\bar{x}_0 = x_0 - x^*. \tag{7}$$

Moreover from (5)

$$d_i^2 = ||x_i - x^*||^2 + 2\bar{\epsilon}_i||x_i - x^*|| + \bar{\epsilon}_i{}^2.$$

From these expressions and the definition of e_i in (6) it follows that

$$e_i = 2(x^* - x_i)'\bar{x}_0 + ||\bar{x}_0||^2 + 2\epsilon_i||\bar{x}_0|| + \eta_i \tag{8}$$

where

$$\eta_i = 2\epsilon_i||x_i - x_0|| + \epsilon_i^2 - 2\bar{\epsilon}_i||x_i - x^*|| - \bar{\epsilon}_i^2 - 2\epsilon_i||\bar{x}_0||.$$

Note that $|||x_i - x_0|| - ||\bar{x}_0||| \leq ||x_i - x^*||$ because of the triangle inequality and the definition of $\bar{x}_0$ in (7). From this and (5) it is easy to see that

$$|\eta_i| \leq (|\epsilon_i| + |\bar{\epsilon}_i|)\gamma_i \tag{9}$$

where $\gamma_i = 2d_i + |\epsilon_i - \bar{\epsilon}_i|$.

3.2 Station Keeping with $n = 3$ Neighbors

In this section we consider the case when $n = 3$. We shall assume that x_1, x_2, and x_3 are not co-linear. Note first that we can write

$$\dot{\bar{x}}_0 = u \tag{10}$$

because of (2) and the fact that $\bar{x}_0 = x_0 - x^*$. Let

$$e = \begin{bmatrix} e_1 - e_3 \\ e_2 - e_3 \end{bmatrix}$$

and define $q = B\bar{x}_0$, where

$$B = 2\left[x_3 - x_1 \; x_3 - x_2\right]'. \tag{11}$$

The error model for this case is then

$$e = q + \epsilon||B^{-1}q|| + \eta \tag{12}$$
$$\dot{q} = Bu \tag{13}$$

where

$$\epsilon = 2\begin{bmatrix}\epsilon_1 - \epsilon_3\\ \epsilon_2 - \epsilon_3\end{bmatrix} \qquad \eta = \begin{bmatrix}\eta_1 - \eta_3\\ \eta_2 - \eta_3\end{bmatrix}.$$

Our assumption that the x_i are not co-linear implies that B is non-singular. Note that since B is nonsingular, $x_0 = x^*$ whenever $q = 0$. This in turn will be the case when $e = 0$ provided $\epsilon = 0$ and $\eta = 0$. The term $||B^{-1}q||\epsilon$ can be regarded as a perturbation and can be dealt with using standard small gain arguments. Essentially linear error models like (12), (13) can also be derived for any $n > 3$.

3.3 Station Keeping with $n = 2$ Neighbors

In the two-neighbor case we've assumed that $|d_1 - d_2| < ||x_1 - x_2|| < d_1 + d_2$ and thus that two solutions x^* to (4) exist. We will assume that $\bar{x}_0$ has been defined so that $||\bar{x}_0(0)||$ is the smaller of the two possibilities. As before, and for the same reason, (10) holds. For this version of the problem we define

$$e = \begin{bmatrix}e_1\\ e_2\end{bmatrix}.$$

Let $q = B\bar{x}_0$, where now

$$B = 2\left[x^* - x_1 \; x^* - x_2\right]'. \tag{14}$$

The error model for this case is then

$$e = q + \epsilon||B^{-1}q|| + ||B^{-1}q||^2\mathbf{1} + \eta \tag{15}$$
$$\dot{q} = Bu \tag{16}$$

where

$$\mathbf{1} = \begin{bmatrix}1\\ 1\end{bmatrix} \qquad \epsilon = 2\begin{bmatrix}\epsilon_1\\ \epsilon_2\end{bmatrix} \qquad \eta = \begin{bmatrix}\eta_1\\ \eta_2\end{bmatrix}.$$

Note that our assumption that $|d_1 - d_2| < ||x_1 - x_2|| < d_1 + d_2$ implies that x_1, x_2, x^* are not co-linear. This in turn implies that B is non-singular. The essential difference between this error model and the error model for the three neighbor case is that the two-neighbor agent model has a quadratic function of state in its readout equation whereas the three-neighbor error model does not.

4 Station Keeping Supervisory Controller

In this section we will develop a set of controller equations aimed at solving the station keeping problem with three neighbors. Because of its properties, the controller we propose can also be used for the two neighbor version of the problem; however in this case meaningful results can only be claimed if agent 0 starts out at a position which is sufficiently close to its target x^*. For ease of reference, we repeat the error equations of interest.

$$e = q + \epsilon ||B^{-1}q|| + \eta \tag{17}$$

$$\dot{q} = Bu. \tag{18}$$

In the sequel we will assume that $||\epsilon|| \leq \epsilon^*, \quad t \geq 0$ where ϵ^* is a positive constant which satisfies the constraint

$$\epsilon^* < \frac{1}{||B^{-1}||}. \tag{19}$$

Note that this constraint says that the allowable measurement error bound will decrease as agents 1, 2, and 3 are positioned closer and closer to co-linear and/or further and further away from agent 0. While we are unable to fully justify this assumption at this time, we suspect that it is intrinsic and is not specific to the particular approach to station keeping which we are following. Our suspicion is prompted in part by the observation that the map $q \longmapsto q + \epsilon ||B^{-1}q||$ will be invertible for all $||\epsilon|| \leq \epsilon^*$ if and only if (19) holds.

The type of control system we intend to develop assumes that B is unknown, but requires one to define at the outset a closed bounded subset of 2×2 non-singular matrices $\mathcal{P} \subset \mathbb{R}^{2\times 2}$ which is big enough so that it can be assumed that $B \in \mathcal{P}$. It is clear that because of the non-singularity requirement, just about any reasonably defined parameter space $\mathcal{P}$ which satisfies these conditions would not be convex, or even the union of a finite number of convex sets. This has important practical implications which we will elaborate on later.

The supervisory control system to be considered consists of a "multi-estimator" $\mathbb{E}$, a "multi-controller" $\mathbb{C}$, a "monitor" $\mathbb{M}$ and a "dwell-time switching logic" $\mathbb{S}$. These terms and definitions have been discussed before in [14, 15] and elsewhere. They are fairly general concepts, have specific meanings, and apply to a broad range of problems. Although there is considerable flexibility in how one might define these component subsystems, in this paper we shall be quite specific. The numbered equations which follow, are the equations which define the supervisory controller we will consider.

4.1 Multi-Estimator $\mathbb{E}$

For the problem of interest, the multi-estimator $\mathbb{E}$ is defined by the two equations

$$\dot{z}_1 = -\lambda z_1 + \lambda e \tag{20}$$

$$\dot{z}_2 = -\lambda z_2 + u \tag{21}$$

where λ is a design constant which must be positive but is otherwise unconstrained.

Note that the signal $\rho = z_1 + Bz_2 - q$ satisfies

$$\dot{\rho} = -\lambda\rho + \lambda(\epsilon||B^{-1}q|| + \eta).$$

For $P \in \mathcal{P}$, let $\bar{e}_P$ denote the Pth *output estimation error*

$$\bar{e}_P = z_1 + Pz_2 - e.$$

The relevant relationships between these signals when $P = B$ can be conveniently described by the block diagram in Figure 1. The diagram describes

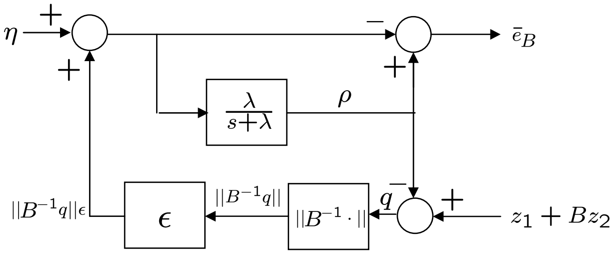

Fig. 1. Subsystem

a nonlinear dynamical system with inputs η and $z_1 + Bz_2$ and outputs $\bar{e}_B$. It is easy to verify that this system is globally exponentially stable with stability margin no smaller than $\lambda(1-\epsilon^*||B^{-1}||)$ because of the measurement constraint (19) discussed earlier. The diagram clearly implies that if ϵ and η were 0, $\bar{e}_B$ would tend to 0; in this case $z_1 + Bz_2$ would therefore be an asymptotically correct estimate of $e = q$. We exploit these observations below.

4.2 Multi-Controller $\mathbb{C}$

The multi-controller $\mathbb{C}$ we propose to study is simply

$$u = -\lambda \widehat{B}^{-1} e \tag{22}$$

where $\widehat{B}$ is a suitably defined piecewise constant switching signal taking values in $\mathcal{P}$. The definition of u has been crafted so that the "closed-loop parameterized system" matrix $-\lambda PP^{-1}$ is stable with "stability margin" λ for all $P \in \mathcal{P}$. Other controllers which accomplish this could also be used {e.g., $u = -\lambda\widehat{B}^{-1}(z_1 + \widehat{B}z_2)$}. The consequence of this definition of u is predicted by the certainty equivalence stabilization theorem [10] and is as follows. Let $\bar{e}_{\widehat{B}} = z_1 + \widehat{B}z_2 - e$ and define the so-called *injected sub-system* to be the system with input $\bar{e}_{\widehat{B}}$ and output $z_1 + Bz_2$ which results when $z_1 + Bz_2 - \bar{e}_{\widehat{B}}$ is substituted for e in the closed loop system determined by (20), (21) and (22). Thus

$$\begin{aligned}\dot{z}_1 &= \lambda\widehat{B}z_2 - \lambda\bar{e}_{\widehat{B}} \\ \dot{z}_2 &= -\lambda\widehat{B}^{-1}z_1 - 2\lambda z_2 + \lambda\widehat{B}^{-1}\bar{e}_{\widehat{B}}.\end{aligned}$$

Certainty equivalence implies that this system, viewed as a dynamical system with input $\bar{e}_{\widehat{B}}$, is also stable with stability margin λ for each fixed $\widehat{B} \in \mathcal{P}$. In this special case one can deduce this directly using the state transformation $\{z_1, z_2\} \longmapsto \{z_1, z_1 + \widehat{B}z_2\}$. For this system to have stability margin λ means that for any positive number $\lambda_0 < \lambda$ the matrix $\lambda_0 I + A(\widehat{B})$ is exponentially stable for all constant $\widehat{B} \in \mathcal{P}$. Here

$$A(\widehat{B}) = \begin{bmatrix} 0 & \lambda\widehat{B} \\ -\lambda\widehat{B}^{-1} & -2\lambda I \end{bmatrix}$$

which is the state coefficient matrix of the injected system.

In the sequel, we fix λ_0 at any positive value such that $\lambda_0 < \lambda(1 - \epsilon^*)||B||^{-1}$. This number turns out to be a lower bound on the convergence rate for the entire closed-loop control system.

We need to pick one more positive design parameter, called a *dwell time* τ_D. This number has to be chosen large enough so that the injected linear system defined above is exponentially stable with stability margin λ for every "admissible" piecewise constant switching signal $\widehat{B} : [0, \infty) \to \mathcal{P}$, where by *admissible* we mean a piecewise constant signal whose switching instants are separated by at least τ_D time units. This is easily accomplished because each $\lambda_0 I + A(P),\ \ P \in \mathcal{P}$ is a stability matrix. All that's required then is to pick τ_D large enough so that the induced norm {any matrix norm} of each matrix $e^{\{\lambda_0 I + A(P)\}t}$, $P \in \mathcal{P}$, is less than 1.

It is useful for analysis to add to Figure 1, two copies of the injected system just defined, one $\{\Sigma_1\}$ with output $e = z_1 + Bz_2 - \bar{e}_{\widehat{B}}$ and the other $\{\Sigma_2\}$ with output $z_1 + Bz_2$. The multiple copies are valid because the injected system is an exponentially stable linear system. The resulting system is shown in Figure 2.

Note that if there were a gain between $\bar{e}_B$ and $\bar{e}_{\widehat{B}}$, and if ϵ were small enough, the overall system shown in Figure 2 would be exponentially stable

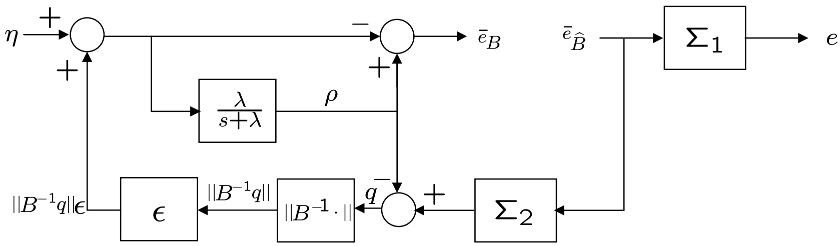

Fig. 2. Subsystem for analysis

and bounded η would produce bounded e. We return to this observation later.

4.3 Monitor $\mathbb{M}$

The state dynamic of monitor $\mathbb{M}$ is defined by the equation

$$\dot{W} = -2\lambda_0 W + \begin{bmatrix} z_1 - e \\ z_2 \end{bmatrix} \begin{bmatrix} z_1 - e \\ z_2 \end{bmatrix}' \tag{23}$$

where W is a "weighting matrix" which takes values in the linear space $\mathcal{X}$ of 4×4 symmetric matrices; although not crucial, for simplicity we will require $\mathbb{M}$ to be initialized at zero; thus $W(0) = 0$. This clearly implies that $W(t)$ is positive semi-definite for all $t \geq 0$. Note that it takes only 10 differential equations rather than 16 to generate W because of symmetry.

The output of $\mathbb{M}$ - First Pass

The output of $\mathbb{M}$ is a parameter dependent "monitoring signal" which for the moment we define to be $\mu_P = M(W, P)$ where $M : \mathcal{X} \times \mathcal{P} \to \mathbb{R}$ is the scalar-valued function

$$M(X, P) = \text{trace}\{\begin{bmatrix} I & P \end{bmatrix} X \begin{bmatrix} I & P \end{bmatrix}'\}.$$

The μ_P are helpful in motivating the definition of $\mathbb{M}$ and the switching logic $\mathbb{S}$ which follows; however, they are actually not used anywhere in the implemented system. It is obvious that they could not be because there are infinitely many of them.

Note that for any $P \in \mathcal{P}$,

$$\dot{\mu}_P = -2\lambda_0 \mu_P + \text{trace}(\{z_1 - e + Pz_2\}\{z_1 - e + Pz_2\}')$$

so

$$\dot{\mu}_P = -2\lambda_0 \mu_P + ||z_1 - e + Pz_2||^2.$$

But $\bar{e}_P = z_1 - e + Pz_2$. Therefore

$$\dot{\mu}_P = -2\lambda_0\mu_P + ||\bar{e}_P||^2$$

and

$$M(W,P) = \int_0^t e^{-2\lambda_0(t-s)}||\bar{e}_P||^2 ds.$$

Thus if we introduce the exponentially weighted 2-norm

$$||\omega||_t = \sqrt{\int_0^t \{e^{\lambda_0 s}||\omega(s)||\}^2 ds}$$

where ω is a piecewise continuous signal, then

$$M(W(t),P) = e^{-2\lambda_0 t}||\bar{e}_P||_t^2, \quad t \geq 0.$$

Minimizing $M(W(t),P)$ with respect to P and setting $\widehat{B}(t)$ to the resulting minimizing value, would then yield an inequality of the form

$$||\bar{e}_{\widehat{B}}||_t \leq ||e_B||_t.$$

Were it possible to accomplish this at every instant of time and were $\widehat{B}$ changing slowly enough so that all of the time-varying subsystems in Figure 2 were exponentially stable, then one could conclude that for ϵ^* sufficiently small, the resulting overall system with input η and output e would be stable with respect to the exponentially weighted norm we've been discussing. It is of course not possible to carry out these steps instantly and even if it were, $\widehat{B}$ would likely be changing too fast for the time-varying subsystems in Figure 2 to be exponentially stable. Were we to continue with this definition of μ_P, we would nonetheless, want to minimize $M(W(t),P)$ from time to time and in doing so would end up with an input-output stable system. In fact the implementation of dwell time switching proposed in [3] requires such minimizations to be carried out. But were we to proceed with this approach, we'd run head on into an important practical problem which we want to address.

A Non-Convex Parameter Space

Note that even though $M(X,P)$ is a quadratic positive semi-definite function of the elements of P, the problem of minimizing $M(X,P)$ over $\mathcal{P}$ is still very complex because $\mathcal{P}$ is not typically convex or even a finite union of convex sets. Thus if we were to use such a parameter space and proceed as we've just outlined, we'd be faced with an intractable non-convex optimization problem. The root of the problem stems from the requirement that the algebraic curve

$$\mathcal{C} = \{P : p_{11}p_{22} - p_{12}p_{21} = 0\}$$

in $\mathbb{R}^{2\times 2}$ on which P is singular cannot intersect $\mathcal{P}$. One way to deal with this difficulty relies on an idea called "cyclic switching" which was specifically devised to deal with this type of problem [17, 13]. Cyclic switching is roughly as follows. First $\mathcal{P}$ is allowed to contain singular matrices, in which case it is reasonable to assume that it is a finite union of compact convex sets. Minimization over $\mathcal{P}$ thus becomes a finite number of standard convex programming problems. For minimizing values of $\widehat{B}$ which turn out to be close to or on $\mathcal{C}$, one uses a specially structured switching controller in place of (22) – one which does not require $\widehat{B}$ to be nonsingular. This controller is used for a specific length of time over which a "switching cycle" takes place. At the end of the cycle, minimization of $M(W, \widehat{B})$ is again carried out; if $\widehat{B}$ is again close to $\mathcal{C}$, another switching cycle is executed. On the other hand, if $\widehat{B}$ is not close to $\mathcal{C}$, the standard certainty equivalence control (22) is used. Cyclic switching is completely systematic and can be shown to solve the singularity problem of interest here. The main disadvantage of cyclic switching is that it introduces additional complexity.

There is another possible way to deal with the singularity problem. What we'd really like is to construct a parameter space $\mathcal{P}$ which is a finite union of convex sets, defined so that every matrix in $\mathcal{P}$ is nonsingular and, in addition, the matrices in $\mathcal{P}$ correspond to a "large" class of possible positions of agents $1, 2, 3$. Keep in mind that the convex subsets whose union defines such a $\mathcal{P}$, *can* overlap. This suggests the following problem.

Problem 1 (Convex Covering Problem). Suppose that we are given a compact subset $\mathcal{P}_0$ of a finite dimensional space which is disjoint from a second closed subset $\mathcal{C}$ {typically an algebraic curve}. Define a *convex cover* of $\mathcal{P}_0$ to mean a finite set of possibly overlapping convex subsets $\mathcal{E}_i$ such that the union of the $\mathcal{E}_i$ contains $\mathcal{P}_0$ but is disjoint from $\mathcal{C}$. One could then define $\mathcal{P}$ to be the union of the $\mathcal{E}_i$.

The existence of such a convex cover can be established as follows[2]. Let d denote the shortest distance between $\mathcal{P}_0$ and $\mathcal{C}$; thus $d = \min\{||p - s|| : p \in \mathcal{P}_0,\ s \in \mathcal{C}\}$. Since $\mathcal{P}_0$ and $\mathcal{C}$ are disjoint, $d > 0$. Let r be any positive number less than d and for each $p \in \mathcal{P}_0$ let $\mathcal{B}(p) = \{q : ||q - p|| < r,\ q \in \mathcal{P}_0\}$. Then for each $p \in \mathcal{P}_0$, the closure of $\mathcal{B}(p)$ and $\mathcal{C}$ are disjoint. Moreover the set of all $\mathcal{B}(p)$ is an open cover of $\mathcal{P}_0$. Thus by the Heine-Borel Theorem, there is a finite subset of the $\mathcal{B}(p)$, say $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_m$ which covers $\mathcal{P}_0$. Setting $\mathcal{E}_i$ equal to the closure of $\mathcal{B}_i$, $i \in \{1, 2, \ldots, m\}$ thus provides a convex cover of $\mathcal{P}_0$ whose union is disjoint from $\mathcal{C}$. Of course this construction would typically produce a cover containing many more convex subsets than might be needed. The question then is how might one going about constructing a convex cover consisting of the smallest number of subsets possible? This unfortunately appears to be a very difficult problem. Nonetheless, its solution could provide an attractive

[2] We thank Ji Liu for pointing this out to us.

alternative to the approach to station keeping which we've outlined in this paper.

There is a third way to avoid the tractability problem which is the approach which we will take here. The key idea is to use a different parameterization which we describe next.

Re-parameterization

Let $\mathcal{U}$ denote the set of all 2×2 matrices U, where each U is a matrix of 0's, 1's and -1's having exactly one nonzero entry in each row and column; there are exactly eight such matrices. It is known [12] that any 2×2 nonsingular matrix M can be written as $M = U(I + L)S$ for some $U \in \mathcal{U}$, some strictly lower triangular matrix L and some symmetric positive definite matrix S. This suggests that we consider a parameter space

$$\mathcal{P} = \{U(I + L)S : \{U, L, S\} \in \mathcal{U} \times \mathcal{L} \times \mathcal{S}\}$$

where $\mathcal{L}$ is a compact, convex subset of the linear space of strictly lower triangular 2×2 matrices and $\mathcal{S}$ a compact, convex subset of the convex set of all 2×2 positive definite matrices. Notice that this definition of $\mathcal{P}$ satisfies both the compactness requirement and the requirement that its elements are all non-singular matrices. Of course one needs to also make sure that $\mathcal{L}$ and $\mathcal{S}$ are large enough so that $B \in \mathcal{P}$. We will say more about how to do this later. For the present we will assume that $B \in \mathcal{P}$ and thus that there are matrices $U_B \in \mathcal{U},\ L_B \in \mathcal{L}$ and $S_B \in \mathcal{S}$ such that

$$B = U_B(I + L_B)S_B.$$

In the sequel we will show that it is possible to meaningfully redefine the type of optimization referred to above as the problem of minimizing a function $J(U, L, S)$ over the set $\mathcal{U} \times \mathcal{L} \times \mathcal{S}$. While this set is not convex, $\mathcal{L} \times \mathcal{S}$ is. Moreover, as we shall see, for each fixed $U \in \mathcal{U},\ J(U, L, S)$ is a convex, quadratic function of the entries in L and S. Because of this, the minimization of $J(U, L, S)$ over $\mathcal{U} \times \mathcal{L} \times \mathcal{S}$ boils down to solving eight convex programming problems, one for each $U \in \mathcal{U}$.

The Output of $\mathbb{M}$ – Second Pass

In the light of the preceding discussion we now re-define $\mathbb{M}$'s output to be $\mu_{\{U,L,S\}} = M(W, U, L, S)$ where now $M : \mathcal{X} \times \mathcal{U} \times \mathcal{L} \times \mathcal{S} \to \mathbb{R}$ is

$$M(X, U, L, S) = \mathrm{trace}\{\left[(I - L)U'\ S\right] X \left[(I - L)U'\ S\right]'\}. \tag{24}$$

In this case it is easy to see that

$$M(W(t), U, L, S) = e^{-2\lambda_0 t}||(I - L)U'\bar{e}_P||_t^2, \quad t \geq 0$$

where $P = U(I + L)S$. In deriving this expression for M we've made use of the easily verified formulas $U' = U^{-1},\ U \in \mathcal{U}$ and $(I + L)^{-1} = I - L,\ L \in \mathcal{L}$.

The matrix $\widehat{B}$ used in the definition of u in (22) is now defined by the formula

$$\widehat{B} = \widehat{U}(I + \widehat{L})\widehat{S} \tag{25}$$

where $\{\widehat{U}, \widehat{L}, \widehat{S}\}$ is a piecewise constant switching signal taking values in $\mathcal{U} \times \mathcal{L} \times \mathcal{S}$. This signal will be generated by a "dwell-time switching logic" which will be described next.

4.4 Dwell-Time Switching Logic $\mathbb{S}$

For our purposes a *dwell-time switching logic* $\mathbb{S}$, is a hybrid dynamical system whose input and output are W and $\widehat{B}$ respectively, and whose state is the ordered triple $\{X, \tau, \{\widehat{U}, \widehat{L}, \widehat{S}\}\}$. Here X is a discrete-time matrix which takes on sampled values of W, and τ is a continuous-time variable called a *timing*

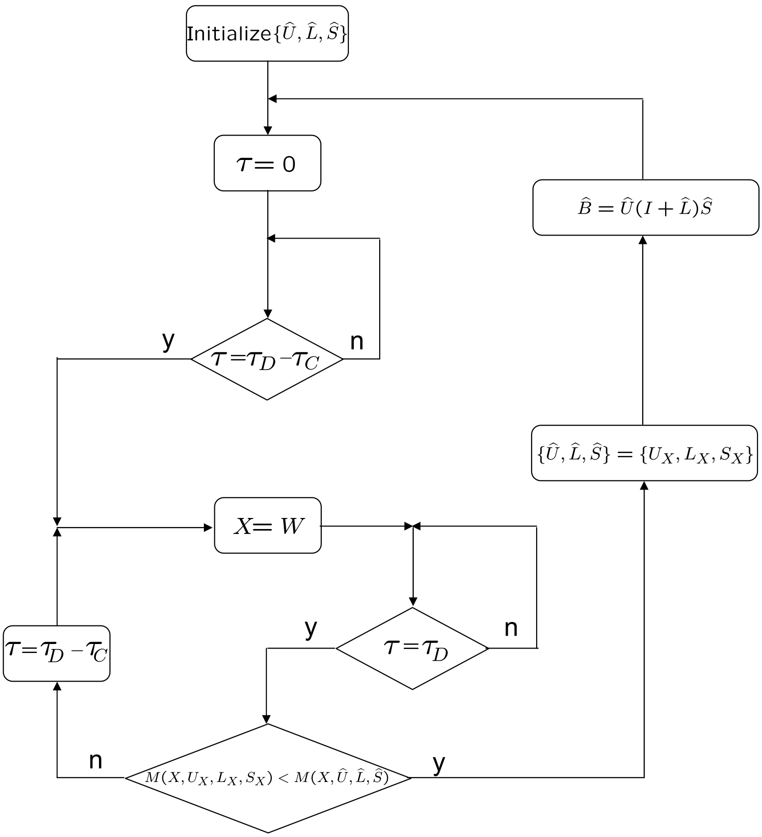

Fig. 3. Dwell-time switching logic $\mathbb{S}$

signal. τ takes values in the closed interval $[0, \tau_D]$. Also assumed pre-specified is a *computation time* $\tau_C \leq \tau_D$ which bounds from above for any $X \in \mathcal{W}$, the time it would take to compute a value $\{U, L, S\} \in \mathcal{U} \times \mathcal{L} \times \mathcal{S}$ which minimizes $M(X, U, L, S)$. Between "event times," τ is generated by a reset integrator according to the rule $\dot{\tau} = 1$. Event times occur when the value of τ reaches either $\tau_D - \tau_C$ or τ_D; at such times τ is reset to either 0 or $\tau_D - \tau_C$ depending on the value of $\mathbb{S}$'s state. $\mathbb{S}$'s internal logic is defined by the flow diagram shown in Figure 3 where $\{U_X, L_X, S_X\}$ denotes a value of $\{U, L, S\} \in \mathcal{U} \times \mathcal{L} \times \mathcal{S}$ which minimizes $M(X, U, L, S)$.

The definition of $\mathbb{S}$ clearly implies that its output $\widehat{B}$ is an admissible switching signal. This means that switching cannot occur infinitely fast and thus that existence and uniqueness of solutions to the differential equations involved is not an issue.

Note that implementation of the switching logic just described requires an algorithm capable of minimizing trace$\{M(X, U, L, S)\}$ over $\mathcal{U} \times \mathcal{L} \times \mathcal{S}$ for various values of $X \in \mathcal{X}$. As we've already explained, for each fixed $U \in \mathcal{U}$, and $X \in \mathcal{X}$, minimization of trace$\{M(X, U, L, S)\}$ reduces to a convex programming problem. Thus for each $X \in \mathcal{X}$, it is enough to solve eight convex programming problems, one for each value of $U \in \mathcal{U}$; the results of these eight computations can then be compared to find the values of U, L and S which attain a global minimum of trace$\{M(X, U, L, S)\}$ over $\mathcal{U} \times \mathcal{L} \times \mathcal{S}$. In other words, by making use of the parameterization we've been discussing, we've been able to reformulate the overall adaptive algorithm in such a way that at each event time all that is necessary is to solve eight, independent quadratic programming problems, one for each $U \in \mathcal{U}$. Of course each of these eight problems may still be challenging. In Section 7 we will explain how each can be reformulated as a semi-definite programming problem.

5 Results

The results which follow rely heavily on the following proposition which characterizes the effect of the monitor-dwell time switching logic subsystem.

Proposition 1. *Suppose that $W(0) = 0$, that $\widehat{B} = \widehat{U}(I + \widehat{L})\widehat{S}$ is the response of the monitor-switching logic subsystem $\{\mathbb{M}, \mathbb{S}\}$ to any continuous input signals e, z_1, and z_2 taking values in $\mathbb{R}^2$, and that for $\{U, L, S\} \in \mathcal{U} \times \mathcal{L} \times \mathcal{S}$, $\bar{e}_P = (z_1 - e) + P z_2$ where $P = U(I + L)S$. For each real number $\gamma > 0$ and each fixed time $T > 0$, there exists piecewise-constant signals $H : [0, \infty) \to \mathbb{R}^{2\times 4}$ and $\psi : [0, \infty) \to \{0, 1\}$ such that*

$$|H(t)| \leq \gamma, \quad t \geq 0 \tag{26}$$

$$\int_0^\infty \psi(t) dt \leq 4(\tau_D + \tau_C) \tag{27}$$

and

$$||(1-\psi)(\bar{e}_{\widehat{B}} - Hz) + \psi \bar{e}_B||_T \leq \delta ||\bar{e}_B||_T \tag{28}$$

where $z = \begin{bmatrix} z_1' & z_2' \end{bmatrix}'$,

$$\delta = 1 + 8\alpha^2 \left(\frac{1 + \text{diameter}\{\mathcal{P}\}}{\gamma} \right)^4,$$

and

$$\alpha = \max_{L \in \mathcal{L}} ||I + L||.$$

This proposition is a minor modification of a similar proposition proved in [14, 15]. The proposition summarizes the key consequences of dwell time switching which are needed to analyze the system under consideration. While the inequality in (28) is more involved than the inequality $||\bar{e}_{\widehat{B}}||_t \leq ||\bar{e}_B||_t$ mentioned earlier, the former is provably correct whereas the latter is not. Despite its complexity, (28) can be used to establish input-output stability with respect to the exponentially weighted norm $|| \cdot ||_t$. The idea is roughly as follows. Fix $T > 0$ and pick γ small enough so that $\lambda_0 I + A(\widehat{B}) + (1-\psi)D(\widehat{B})H$ is exponentially stable where $A(\widehat{B})$ is the state evolution matrix of the injected system defined at the beginning of Section 4.2 and $D(\widehat{B}) = \begin{bmatrix} -\lambda I' & \lambda(\widehat{B}^{-1})' \end{bmatrix}'$. The fact that ψ has a finite $\mathcal{L}^1$ norm {cf. (27)}, implies that $\lambda_0 I + A(\widehat{B}) + (1-\psi)D(\widehat{B})H + \psi \begin{bmatrix} 0 & \widehat{B} - B \end{bmatrix}$ is exponentially stable as well. Next define

$$\bar{e} = (1-\psi)(\bar{e}_{\widehat{B}} - Hz) + \psi \bar{e}_B.$$

Then

$$||\bar{e}||_T \leq \delta ||\bar{e}_B||_T \tag{29}$$

because of (28). The definition of $\bar{e}$ implies that

$$\bar{e}_{\widehat{B}} = \bar{e} + (1-\psi)Hz + \psi \begin{bmatrix} 0 & \widehat{B} - B \end{bmatrix} z.$$

Substitution into the injected system defined earlier yields the exponentially stable system

$$\dot{z} = \{A(\widehat{B}) + (1-\psi)D(\widehat{B})H + \psi \begin{bmatrix} 0 & \widehat{B} - B \end{bmatrix}\}z + D(\widehat{B})\bar{e}$$

with input $\bar{e}$. Now add to Figure 1, two copies of the system just defined, one $\{\bar{\Sigma}_1\}$ with output $e = \begin{bmatrix} I & \widehat{B} \end{bmatrix} z - \{\bar{e} + (1-\psi)Hz + \psi \begin{bmatrix} 0 & \widehat{B} - B \end{bmatrix} z\}$ and the other $\{\bar{\Sigma}_2\}$ with output $z_1 + Bz_2 = \begin{bmatrix} I & B \end{bmatrix} z$. Like before, the multiple copies are valid because the matrix $A(\widehat{B}) + (1-\psi)D(\widehat{B})H + \psi \begin{bmatrix} 0 & \widehat{B} - B \end{bmatrix}$ is exponentially stable. The resulting overall system is shown in Figure 4.

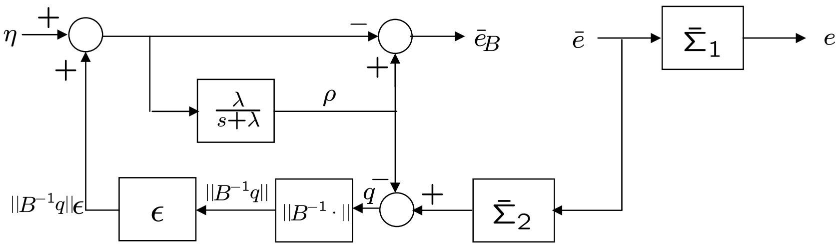

Fig. 4. Snapshot at time T of the overall subsystem for analysis

In the light of (29) it is easy to see that if the bound ϵ^* on ϵ is sufficiently small, the induced gain of this system from η to e with respect to $||\cdot||_T$ is bounded by a finite constant g_T. It can be shown that g_T in turn, is bounded above by a constant g not depending on T [15]. Since this is true for all T, it must be true that g bounds the induced gain from η to e with respect to $||\cdot||_\infty$.

The following results are fairly straightforward consequences of these ideas. Detailed proofs, specific to the problem at hand, can be found in the full-length version of this paper. The results are as follows:

1) If all measurement errors ϵ_i and all miss-alignment errors $\bar{e}_i$ are zero, then, no matter what its initial value, $x_0(t)$ tends to the unique solution x^* to (4) as fast as $e^{-\lambda_0 t}$.
2) If the measurement errors ϵ_i and the miss-alignment errors $\bar{e}_i$ are not all zero, and the ϵ_i sufficiently small, then no matter what its initial value, $x_0(t)$ tends to a value for which the norm of the error e is bounded by a constant times the sum of the norms of the ϵ_i and the $\bar{e}_i$.

Before leaving this section, it should be mentioned that success with the new parameterization we've proposed, of course comes with a price. Note that the gain δ which appears in the statement of Proposition 1 is an increasing function of α, and moreover $\alpha > 1$. Thus the effect of re-parameterization is, in essence, to increase the "gain" around the loop containing $\bar{\Sigma}_2$ in Figure 4. This in turn, reduces the stability margin associated with ϵ and also increases overall induced gain from η to e.

6 Definitions for $\mathcal{L}$ and $\mathcal{S}$

So far we have assumed that $\mathcal{L}$ is a compact, convex subset of the linear space of strictly lower triangular 2×2 matrices and that $\mathcal{S}$ is a compact, convex subset of the set of positive definite 2×2 matrices. The assumptions are sufficient to ensure that any matrix in

$$\mathcal{P} = \{U(I + L)S : (U, L, S) \in \mathcal{U} \times \mathcal{L} \times \mathcal{S}\}$$

is invertible and also that the minimization of

$$M(X, U, L, S) = \text{trace}\{\left[(I - L)U' \ S\right] X \left[(I - L)U' \ S\right]'\}$$

over $\mathcal{L} \times \mathcal{S}$ for any fixed $U \in \mathcal{U}$ and any fixed positive semi-definite 2×2 matrix X, is a convex programming problem. But we've not yet explained how to explicitly define $\mathcal{L}$ and $\mathcal{S}$. To do this, it makes sense to first define bounds for B which are meaningful for the problem at hand. Towards this end, suppose that agent 0 has a limited sensing radius ρ. Since we've assumed that agent 0 can sense the distances to agents $1, 2$, and 3, it must be true that $||x_3 - x_1|| \leq 2\rho$ and $||x_3 - x_1|| \leq 2\rho$. But $B = 2\left[x_3 - x_1 \ x_3 - x_2\right]'$. Prompted by this we will assume that $\sqrt{B'B} \leq \beta_2 I$ where $\beta_2 = 4\rho$.

We've also assumed that agents $1, 2$ and 3 are not positioned along a line; this is equivalent to B being nonsingular. One measure of B's non-singularity, is its smallest singular value. Prompted by this, we will assume that there is a positive number β_1 such that $\sqrt{B'B} \geq \beta_1 I$; β_1 might be chosen empirically to reflect the degree to which the three leader agents are non co-linear in a given formation. We shall assume that such a number has been chosen and moreover that $\beta_1 < \beta_2$. In summary we suppose that bounds β_1 and β_2 have been derived such that

$$\beta_1 I \leq \sqrt{B'B} \leq \beta_2 I \tag{30}$$

where β_1 and β_2 are distinct positive numbers. It is obvious that the set of matrices B satisfying these inequalities is not convex.

Our next objective is to define $\mathcal{L}$ and $\mathcal{S}$ so that any matrix B satisfying (30) is in $\mathcal{P}$. Let $\mathcal{L}$ be the set of all strictly lower triangular 2×2 matrices $L = \left[l_{ij}\right]$ for which

$$|l_{21}| \leq 1 + \sqrt{2}\frac{\beta_2}{\beta_1}. \tag{31}$$

In addition, let $\mathcal{S}$ be the set of all 2×2, symmetric matrices satisfying

$$\sigma_1 I \leq S \leq \sigma_2 I \tag{32}$$

where

$$\sigma_1 = \frac{1}{\left(2\sqrt{1 + \left(\frac{\beta_2}{\beta_1}\right)^2}\right)}\beta_1$$

$$\sigma_2 = \left(2\sqrt{1 + \left(\frac{\beta_2}{\beta_1}\right)^2}\right)\beta_2. \tag{33}$$

It is now shown that any matrix B satisfying (30) is in $\mathcal{P}$.

As a first step, let us note that b_{11} and b_{21} cannot both be zero because B is nonsingular. If $|b_{11}| \geq |b_{21}|$, let

$$U = \begin{bmatrix} \mathrm{sign}\{b_{11}\} & 0 \\ \\ 0 & \mathrm{sign}\{b_{11}d\} \end{bmatrix}$$

$$L = \begin{bmatrix} 0 & 0 \\ \\ \frac{u_{22}b_{21}-u_{11}b_{12}}{|b_{11}|} & 0 \end{bmatrix}$$

$$S = \begin{bmatrix} |b_{11}| & u_{11}b_{12} \\ \\ u_{11}b_{12} & \frac{b_{12}^2+|d|}{|b_{11}|} \end{bmatrix}. \tag{34}$$

On the other hand, if $|b_{21}| > |b_{12}|$, let

$$U = \begin{bmatrix} 0 & -\mathrm{sign}\{b_{21}d\} \\ \\ \mathrm{sign}\{b_{21}\} & 0 \end{bmatrix}$$

$$L = \begin{bmatrix} 0 & 0 \\ \\ \frac{u_{12}b_{11}-u_{21}b_{22}}{|b_{21}|} & 0 \end{bmatrix}$$

$$S = \begin{bmatrix} |b_{21}| & u_{21}b_{22} \\ \\ u_{21}b_{22} & \frac{b_{22}^2+|d|}{|b_{21}|} \end{bmatrix}. \tag{35}$$

In either case it is easy to verify that $B = U(I + L)S$. It is also clear that in either case $U \in \mathcal{U}$, that L is strictly lower triangular and that S is symmetric. Thus to prove that $B \in \mathcal{P}$, it is sufficient to show that in either of the two cases, L and S satisfy (31) and (32) respectively. We will do this only for the case $|b_{11}| \geq |b_{21}|$ as similar reasoning applies to the case $|b_{21}| < |b_{11}|$.

Let us note from (34) that $|l_{21}| \leq \left|\frac{b_{21}}{b_{11}}\right| + \left|\frac{b_{12}}{b_{11}}\right|$. By assumption $|b_{11}| \geq |b_{21}|$; this implies that $\left|\frac{b_{21}}{b_{11}}\right| \leq 1$ so $|l_{21}| \leq 1 + \left|\frac{b_{12}}{b_{11}}\right|$. Now from (30), $\beta_1 \leq \sqrt{b_{11}^2 + b_{21}^2}$, so $\beta_1 \leq \sqrt{2b_{11}^2} = \sqrt{2}|b_{11}|$; also from (30), $|b_{12}| \leq \beta_2$. Therefore $\left|\frac{b_{12}}{b_{11}}\right| \leq \sqrt{2}\frac{\beta_2}{\beta_1}$. It follows that l_{21} satisfies (31).

Next observe that $B'B = S(I+L)'U'U(I+L)S = S(I+L)'(I+L)S$. Now $(I+L)'(I+L) \leq (2+|l_{12}|^2)I$. Therefore $B'B \leq (2+|l_{12}|^2)S^2$. From this and (30), it follows that $S^2 \geq \frac{\beta_1^2}{2+|l_{12}|^2}I$. From (31),

$$l_{21}^2 \leq 2\left(1 + 2\frac{\beta_2^2}{\beta_1^2}\right). \tag{36}$$

Therefore $S^2 \geq \frac{\beta_1^4}{4(\beta_1^2+\beta_2^2)}I = \sigma_1^2 I$.

Finally observe that $S = (I - L)U'B$ and thus that $S^2 = B'U(I - L)'(I - L)U'B$. But $(I-L)'(I-L) \leq (2+|l_{12}|^2)I$. Therefore $S^2 \leq (2+|l_{12}|^2)B'UU'B =$

$(2+|l_{12}|^2)B'B$. From this (30), and (36) it follows that $S^2 \leq 4(1+\frac{\beta_2^2}{\beta_1^2})\beta_2^2 I$. Therefore S satisfies both inequalities in (32). This means that $B \in \mathcal{P}$.

7 Semi-Definite Programming Formulation

Fix $U \in \mathcal{U}$, and let $X \in \mathcal{X}$ be a given positive semi-definite matrix. To implement the dwell time switching logic defined in Section 4.4, it is necessary to make use of an algorithm capable of minimizing over $\mathcal{L} \times \mathcal{S}$, a cost function of the form

$$N(L,S) = \text{trace}\{\left[(I-L)U'\ S\right] X \left[(I-L)U'\ S\right]'\}. \tag{37}$$

Our aim is to explain how to reformulate this convex optimization problem as a convex semi-definite programming problem over the space $\mathcal{Y} \times \mathcal{L} \times \mathcal{Y}$ where $\mathcal{Y}$ is the linear space of 2×2 symmetric matrices[3]. As a first step towards this end, we exploit two easily proved facts. First, if (L_1, S_1) minimizes $N(L,S)$ over $\mathcal{L} \times \mathcal{S}$, then $(\{\left[(I-L_1)U_1'\ S_1\right] X \left[(I-L_1)U_1'\ S_1\right]'\}, L_1, S_1)$ minimizes

$$\bar{N}(Y,L,S) = \text{trace}\{Y\}$$

over $\mathcal{Y} \times \mathcal{L} \times \mathcal{S}$ subject to the constraint that $Y - [(I-L_1)U_1'\ S_1]X[(I-L_1)U_1'\ S_1]'$ is positive semi-definite. Second, if (Y_2, L_2, S_2) minimizes $\bar{N}(Y,L,S)$ over $\mathcal{Y} \times \mathcal{L} \times \mathcal{S}$ subject to the constraint that $Y - [(I-L_1)U_1'\ S_1]X[(I-L_1)U_1'\ S_1]'$ is positive semi-definite, then (L_2, S_2) minimizes $N(L,S)$ over $\mathcal{L} \times \mathcal{S}$. In other words, the optimization problem of interest is equivalent to minimizing the cost $\bar{N}(Y,L,S)$ over $\mathcal{Y} \times \mathcal{L} \times \mathcal{S}$ subject to the constraint

$$Y - \left[(I-L)U'\ S\right] X \left[(I-L)U'\ S\right]' \geq 0. \tag{38}$$

To proceed, let us next observe that the matrix to the left in the above inequality, is the Schur complement of the matrix

$$Q = \begin{bmatrix} I & R'\left[(I-L)U'\ S\right]' \\ \left[(I-L)U'\ S\right]R & Y \end{bmatrix}$$

where R is any matrix such that $X = RR'$. Thus the matrix inequality in (38) is equivalent to the matrix inequality

$$Q \geq 0. \tag{39}$$

Moreover the constraint that $S \in \mathcal{S}$ is equivalent to $S \in \mathcal{Y}$ and the pair of linear matrix inequality constraints $\sigma_2 I - S \geq 0$ and $S - \sigma_1 I \geq 0$. These constraints can be combined with (39) to give finally the constraint

[3] We are indebted to Ali Jadbabai for making us aware of this simplification.

$$\begin{bmatrix} Q & 0 & 0 \\ 0 & \sigma_2 I - S & 0 \\ 0 & 0 & S - \sigma_1 I \end{bmatrix} \geq 0. \tag{40}$$

Thus we've reduced the optimization problem of interest to minimizing $\bar{N}(Y, L, S)$ over $\mathcal{Y} \times \mathcal{L} \times \mathcal{Y}$ subject to (40). Since (31) is equivalent to two linear inequality constraints, the problem to which we've been led is a conventional convex, semi-definite programming problem [22]. Of course to carry out this optimization, one needs also an standard algorithm to factor a positive semi-definite matrix X as $X = RR'$.

8 Concluding Remarks

In this paper we have devised a tractable solution to the three neighbor station keeping problem in which range measurements are the only sensed signals upon which station keeping is to be based. The solution is the same as that in [3] except that here a special parameterization is used to avoid the non-convex optimization problem which must be solved in order to implement the algorithm in [3]. The solution in this paper is provably correct and the performance of the resulting system degrades gracefully in the face of increasing measurement and miss-alignment errors, provided the measurement errors are not too large. We have used standard constructions from adaptive control to accomplish this. Because of the exponential stability of the overall system, the same control algorithm will solve the two agent station keeping problem provided the agent is initially not too far from its target position.

Acknowledgments

This research was supported by the National Science Foundation, the US Army Research Office, and by a gift from the Xerox Corporation.

References

1. L. Armesto and J. Tornero. SLAM based on Kalman filter for multi-rate fusion of laser and encoder measurements. In *Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1860–1865, 2004.
2. M. Cao and A.S. Morse. Mainaining an autonomous agent's position in a moving formation with range-only measurements. In *Proc. of the 2007 European Control Conference*, 2007. to be presented.
3. M. Cao and A.S. Morse. Station-keeping in the plane with range-only measurements. In *Proc. of the 2007 Amer. Contr. Conf.*, 2007. to be presented.

4. S. H. Dandach, B. Fidan, S. Dasgupta, and B.D.O. Anderson. Adaptive source localization by mobile agents. In *Proc. of the 45rd IEEE Conf. on Decision and Contr.*, pages 2045–2050, 2006.
5. S. H. Dandach, B. Fidan, S. Dasgupta, and B.D.O. Anderson. A continuous time linear adaptive source localization algorithm robust to persistent drift. *Systems & Control Letters*, 2007. To appear.
6. M. W. Dissanayake, P. Newman, H. F. Durrant-Whyte, S. Clark, and M. Csorba. An experimental and thoretical investigation into simultaneous localisation and map building. *Experimental Robotics*, IV:265–274, 2000.
7. H.J.S. Feder, J.J. Leonard, and C.M. Smith. Adaptive mobile robot navigation and mapping. *International Journal of Robotics Research*, 18:650–668, 1999.
8. D. Fox, W. Burgard, F. Dellaert, and S. Thrun. Monte Carlo localization: Efficient position estimation for mobile robots. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, 1999.
9. B. Grocholsky, E. Stump, P. Shiroma, and V. Kumar. Control for localization of targets using range-only sensors. In *Proceedings of the International Symposium on Experimental Robotics*, 2006.
10. J.P. Hespanha and A.S. Morse. Certainty equivalence implies detectability. *Systems & Control Letters*, pages 1–13, 1999.
11. J.J. Leonard and H.F. Durrant-Whyte. Simultaneous map building and localization for an autonomous mobile robot. In *Proceedings of the 1991 IEEE/RSJ International Workshop on Intelligent Robots and Systems*, pages 1442–1447, 1991.
12. A.S. Morse. A gain matrix decomposition and some of its applications. *SCL*, 21:1–10, 1993.
13. A.S. Morse. Cyclic switching and supervisory control. In *Proceedings of the 1995 IFAC Symposium on Nonlinear Control Systems Design*, pages 499–502, 1995.
14. A.S. Morse. Analysis of a supervised set-point control system containing a compact continuum of finite dimensional linear controllers. In *Proc. of Mathematical Theory of Networks and Systems*, 2004.
15. A.S. Morse. Logically switched dynamical systems. In *Nonlinear and Optimal Control Theory*, pages 1–84. Springer-Verlag, 2005. To appear.
16. P. Newman. *On the Structure and Solution of the Simultaneous Localisation and Map Building Problem.* Ph.D dissertation, University of Sydney, Australian Centre for Field Robotics, 1999.
17. F.M. Pait and A.S. Morse. A cyclic switching strategy for parameter-adaptive control. *IEEE Trans. on Automat. Contr.*, 39(6):1172–1183, 1994.
18. A.H. Sayed and A. Tarighat. Network-based wireless location. *IEEE Signal Processing Magazine*, pages 24–40, July 2005.
19. R. Smith, M. Self, and P. Cheeseman. Estimating uncertain spatial relationships in robotics. In I. Cox and G. Wilfong, editors, *Autonomous Robot Vehicles.* Springer-Verlag, 1990.
20. E. Stump, B. Grocholsky, and V. Kumar. Extensive representaions and algorithms for nonlinear filtering and estimation. In *Proceedings of the 7th Annual Workshop on the Algorithmic Foundations of Robotics*, 2006.
21. S. Thrun, D. Fox, and W. Burgard. A probabilistic approach to concurrent mapping and localization for mobile robots. *Machine Learning*, 31:29–53, 1998.
22. L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38:49–95, 1996.

23. S.B. Williams, M.W.. Dissanayake, and H.F. Durrant-Whyte. Towards multi-vehicle simultaneous localisation and mapping. In *Proceedings of the 2002 IEEE International Conference on Robotics and Automation*, pages 2743–2748, 2002.
24. D.C.K. Yuen and B.A. MacDonald. An evaluation of the Sequential Monte Carlo technique for simultaneous localisation and map-building. In *Proceedings of the 2003 IEEE International Conference on Robotics and Automation*, pages 1564–1569, 2003.